



Université Claude Bernard



Lyon 1

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **17 décembre 2021**

Nom de famille et prénom de l'auteur : **Madame SIMAO Maryanna Cristiano**

Titre de la thèse : *Transposable elements (TEs) exaptation in Drosophila mojavensis and expression bias of genes and TEs in gonads of D. mojavensis and D. arizonae and their hybrids.*

Résumé



Transposable elements (TEs) are repetitive DNA sequences able to mobilize within and between genomes that due to this ability, besides deleterious effects, can bring advantages to the host genome, such as increasing the protein repertoire or influencing gene regulation. The study material of this thesis was the genome of *Drosophila mojavensis*, a cactophilic species, native to arid and semi-arid regions, which uses as feeding and breeding site that is less rich in resource than those used by most *Drosophila* species and, consequently, such habits must be related to specific adaptations. The biological and ecological characteristics of the species make this genome a good option to analyze the content and characteristics of TEs, as well as to investigate the association of TEs with protein-coding genes (PCG) and long noncoding RNA (lncRNA) genes. Furthermore, even though most genes are common to male and female genomes, genes and TEs may be differentially expressed between testes and ovaries. Analysis of the gonad transcriptomes makes it possible to investigate how or if the expression bias of genes and TEs affects the reproduction of interspecific hybrids. In order to better understand these questions, the expression bias of the gonads of *D. mojavensis* and *D. arizonae* and their reciprocal hybrids was studied in this thesis. In this study, we show that TE content represents 13.27% of the *D. mojavensis* genome and that the proportion of genes harboring TE sequences within exons is 5.9%, and that the proportion of TE orders within these exons differs from that of the overall genome (LINEs: 78.8% > LTR: 18.5% > TIRs: 2.4% > Helitrons: 0.3%). Among all TE insertions found in PCG

and lncRNA genes, only 19 sequences showed preserved protein domains. Several candidate genes for TEs exaptation were identified in the *D. mojavensis* genome and most of them harbor DNA-binding sites, particularly zinc-finger domains, highlighting the role of TEs as a source of functional novelties to genomes, evidencing that this association between specific DNA-binding sites and TEs is recurrent and perhaps one of the major sources of genomic novelties. Most sex-biased genes and TEs are present in both parental species, and there are three times more testis-biased genes and TEs than ovary-biased, highlighting that expression between gonads in *Drosophila* is conserved and characteristic of the group. However, there are specific differences in sex-bias expression between parental species and between hybrids, which may play an important role in reproduction. Furthermore, testis-biased TE families are more diverse than testes-biased genes, this fact may be directly related to the influence of piRNAs on the regulation, mainly due to the fact that piRNAs from the hybrids have a different pattern than those found in the parental species.

Keywords: exaptation, zf-BED domains, lncRNA genes, Sex-biased genes, gonadal gene expression, differentially expressed genes, reproductive isolation, *Drosophila mojavensis*, *Drosophila arizonae*.